

AMENDMENTS TO THE CLAIMS:

1-48 (Cancelled)

49. (New) A machine implemented method for providing statistics characterizing content for language translation, comprising the steps of:

accessing content in a first language, including at least one of web content, content having markup tags therein, and content retrieved by following a link contained in the web content or content having markup tags;

parsing the content into one or more translatable components; and

generating statistics based on the one or more translatable components.

50. (New) The method according to claim 49, further comprising using the statistics to size the content for language translation.

51. (New) The method according to claim 50, wherein the language translation includes human translating the one or more translatable components.

52. (New) The method according to claim 49, further comprising the step of adding the one or more translatable components to a translation list for translation into a second language.

53. (New) The method according to claim 52, wherein:

the first language includes one of English, French, Spanish, German, Portuguese, Italian, Chinese, Korean, and Arabic;

the second language includes one of English, French, Spanish, German, Portuguese, Italian, Japanese, Chinese, Korean, and Arabic; and

the second language is different from the first language.

54. (New) The method according to claim 49, wherein the step of accessing includes retrieving the content in the first language from a source in a first language.

55. (New) The method according to claim 49, wherein each of the translatable components is one of:

a text segment;

an image file;

an audio clip;

a video clip;

a file; and

any combination thereof in an electronic data stream.

56. (New) The method according to claim 49, further comprising the step of generating an identifier for each of the translatable components so that each of the translatable components is accessible via a corresponding identifier.

57. (New) The method according to claim 56, wherein the identifier for a text segment is generated using at least one of a hash code, a checksum, and a mathematical algorithm based on one or more text segments.

58. (New) The method according to claim 49, wherein the statistics includes at least one of a file count, a page count, a text segment count, a unique text segment count, a word count, and a unique word count.

59. (New) The method according to claim 49, wherein the step of parsing is performed based on at least one markup tag contained in the content in the first language.

60. (New) A machine implemented method for providing statistics characterizing content, comprising the steps of:

accessing content including at least one of web content, content having markup tags therein, and content retrieved by following a link contained in the web content or content having markup tags;

parsing the content into one or more components; and

generating statistics based on the one or more components.

61. (New) A system for providing statistics characterizing content for language translation, comprising:

a connection configured for accessing content in a first language, including at least one of web content, content having markup tags therein, and content retrieved by following a link contained in the web content or content having markup tags;

an information processing portion configured for parsing the content into one or more translatable components; and

a statistics generation portion configured for generating statistics based on the one or more translatable components.

62. (New) The system according to claim 61, further comprising a content sizing portion configured for using the statistics to size the content for language translation.

63. (New) The system according to claim 62, wherein the language translation includes human translating the one or more translatable components.

64. (New) The system according to claim 61, wherein the connection is further configured for facilitating retrieval of content from a data source in the first language.

65. (New) The system according to claim 61, wherein each of the translatable components is one of:

- a text segment;
- an image file;
- an audio clip;
- a video clip;
- a file; and
- any combination thereof in an electronic data stream.

66. (New) The system according to claim 61, wherein the statistics includes at least one of a file count, a page count, a text segment count, a unique text segment count, a word count, and a unique word count.

67. (New) The system according to claim 61, wherein the first language includes one of English, French, Spanish, German, Portuguese, Italian, Chinese, Korean, and Arabic.

68. (New) The system according to claim 61, wherein the information processing portion parses the content based on at least one markup tag contained in the content.

69. (New) The system according to claim 61, further comprising a translation list to which the one or more components are added for translation.

70. (New) A machine readable medium having data stored thereon, the data, when read, causing the machine to perform the following:

accessing content in a first language, including at least one of web content, content having markup tags therein, and content retrieved by following a link contained in the web content or content having markup tags;

parsing the content into one or more translatable components; and

generating statistics based on the one or more translatable components.

71. (New) The medium according to claim 70, the data, when read, further causing the machine to use the statistics to size the content for language translation.

72. (New) The medium according to claim 71, wherein the language translation includes human translating the one or more translatable components.

73. (New) The medium according to claim 70, wherein the statistics includes at least one of a file count, a page count, a text segment count, a unique text segment count, a word count, and a unique word count.

74. (New) A machine implemented method for providing statistics characterizing content for language translation, comprising the steps of:

accessing content in a first language, including at least one of web content, content having markup tags therein, and content retrieved by following a link contained in the web content or content having markup tags;

parsing the content into one or more translatable components based on markup tags contained therein; and

generating statistics based on the one or more translatable components, wherein the translatable components are to be human translated.

75. (New) The method according to claim 74, further comprising the steps of:
obtaining tracking information corresponding to a source from where the content is
retrieved; and
utilizing the tracking information when additional content is to be retrieved from the
source.

76. (New) The method according to claim 75, wherein the tracking information includes
at least one of state information or session information.

77. (New) The method according to claim 76, wherein the tracking information is
obtained via a cookie.

78. (New) The method according to claim 74, wherein when the content contains an
HTML form, the method further comprising the steps of:
populating the HTML form automatically with pre-defined information; and
submitting the populated HTML form to the source.

79. (New) The method according to claim 74, wherein the source from where the content
is retrieved is a web site.

80. (New) The method according to claim 74, further comprising:
identifying an image file associated with the content in the first language as a translatable
component;
generating statistics for the image file.

81. (New) The method according to claim 80, wherein the statistics for the image file is
generated manually.

82. (New) The method according to claim 80, wherein the statistics for the image file is generated automatically.